Improving subband spectral estimation using modified AR model

D. Bonacci[a,*], C. Mailhes[a,b]

[a]*TeSA Laboratory, 14-16 Port Saint Etienne, 31000 Toulouse, France*[1]
[b]*ENSEEIHT/IRIT, 2 Rue Camichel, BP 7122, 31071 Toulouse Cedex 7, France*[2]

Abstract

It has already been shown that spectral estimation can be improved when applied to subband outputs of an adapted filterbank rather than to the original fullband signal. In the present paper, this procedure is applied jointly to a novel predictive autoregressive (AR) model. The model exploits time-shifting and is therefore referred to as time-shift AR (TS-AR) model. Estimators are proposed for the unknown TS-AR parameters and the spectrum of the observed signal. The TS-AR model yields improved spectrum estimation by taking advantage of the correlation between subseries that arises after decimation. Simulation results on signals with continuous and line spectra that demonstrate the performance of the proposed method are provided.

Keywords: Subband decomposition; AR modeling; Spectral estimation

1. Introduction

Spectral estimation has been for a long time one of the main field of interests in signal processing theory and applications. During the past 10 years, subband decomposition has been shown to be a powerful tool for spectral analysis and data coding. Indeed, past and recent audio-coding standards are based on subband decomposition [1,2]. Among all existing standards, MP3 may be the most famous one but more recent works, such as [3,4] can also be of interest. Within the field of spectral analysis, and more precisely of parametric modeling, benefits of subband decomposition for order selection are illustrated in [5] in the case of two separate narrow peaks. Although some authors use subband decomposition to improve classical spectral estimation (based on the Fourier transform) [6], subband decomposition is more efficient when it is applied in combination with parametric spectral estimation methods [7,8]. In these papers, subband spectral estimation is shown to yield better performance than applying spectral estimation on the original fullband process. This has been shown for a bank of ideal infinitely sharp bandpass filters. However, some experimental results have highlighted that the improvements brought by subband spectral

*Corresponding author. Tel.: +33 5 61 24 73 81; fax: +33 5 61 24 73 73.
E-mail addresses: david.bonacci@enseeiht.fr (D. Bonacci), corinne.mailhes@enseeiht.fr (C. Mailhes).
URLs: http://dabo1.free.fr/, http://www.enseeiht.fr/~mailhes/.
[1]Telecommunications for Space and Aeronautics.
[2]National Polytechnic Institute of Toulouse.

estimation still remain in the case of non-ideal filterbanks such as modified quadrature-mirror filters (QMFs) or cosine modulated filterbanks [8].

Thus, the performance of traditional or parametric spectral estimation methods can improve when applied to signals filtered by an appropriate filterbank rather than applied to the corresponding fullband signal. Some of the benefits provided by subband decomposition can be given as follows:

(B_1) Model order reduction and consequently condition number decreasing for autocorrelation matrices [9].
(B_2) Frequency spacing and local Signal to Noise Ratio (SNR) increase by the decimation ratio (for signals composed by a sum of sinusoids corrupted by additive noise) [10].
(B_3) Whitening of noise in the subbands [10].
(B_4) Linear prediction error power reduction for AR estimation [7].

Obviously, subband spectral estimation has also some drawbacks, mainly

(D_1) Spectral overlapping (aliasing): when using non-ideal filterbanks, the same harmonic component may appear in two contiguous subbands at two different frequencies.
(D_2) Relative variance increase for autocorrelation estimators (due to decimation).

The first drawback (D_1) has already been addressed in two recent papers [11,12]. All simulation results of the present paper have been conducted using the procedure described in [12] using an uniform filterbank with M subbands and a decimation ratio of M as shown in Fig. 1. The main subject of this paper is to tackle the second drawback (D_2) of subband spectral estimation. In order to decrease

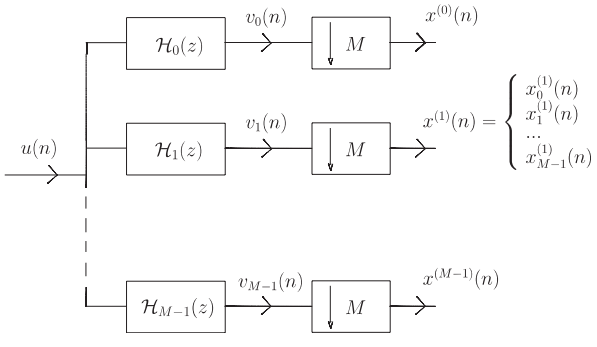

Fig. 1. Uniform filterbank with M subbands.

spectral estimation variance, a new subband spectral estimator is proposed. This method, based on AR modeling, exploits the existence of M different subseries available after decimation on each subband.

The paper is organized as follows. Section 2 states the problem of subband decomposition. The proposed model is defined in Section 3. Estimation of the model parameters and properties of the model error are outlined in Section 4. Section 5 derives the model power spectral density (PSD) and Section 6 discusses the benefits of this new method. Practical implementation and performances of the proposed model are highlighted in Section 7. Concluding comments are given in Section 8.

2. Subband decomposition

Let us consider one branch of the analysis filterbank of Fig. 1. For sake of simplicity, the subband's subscript is omitted. Let $\{u(n), n \in \mathbb{Z}\}$ denote the input signal to be analyzed. This input signal is assumed to be a wide stationary random sequence. It is first filtered by a bandpass filter with a transfer function $\mathscr{H}(z)$ to isolate the frequency band of interest. The resulting bandpass signal $\{v(n), n \in \mathbb{Z}\}$ is then directly reduced in sampling rate by an M-sample compressor giving the final output $\{x(n), n \in \mathbb{Z}\}$.

This scheme is common in subband processing. However, the main idea of the present paper is to take advantage of a specific way of down-sampling. Indeed, depending on the decimator initialization, M different realizations of the random sequence $\{x(n), n \in \mathbb{Z}\}$ can be obtained. These realizations are denoted $x_0(n), \dots, x_{M-1}(n)$ in what follows and can be expressed as

$$x_m(n) = v(Mn - m), \quad m = 0, \dots, M-1. \tag{1}$$

These M subseries can be viewed as different sampling results of the same random sequence $\{x(n), n \in \mathbb{Z}\}$. Therefore, all these subseries $\{x_m(n)\}$ have the same autocorrelation function and the same PSD.

3. Time-shift AR (TS-AR) model definition

Usually, when parametric spectral estimation is performed in a subband decomposition context, parametric model estimation is applied on one of the subseries of (1). Among all parametric models, AR-based ones are the most widely used. These

models may have various forms but all require autocorrelation estimation. Therefore, when applying such methods directly on subseries of the form of (1), the performance of spectral estimation depend on the estimation of the autocorrelation function:

$$r_x(k) = E[x(n)x^*(n-k)] = r_v(Mk), \quad k \in \mathbb{Z}, \tag{2}$$

where $E[.]$ denotes mathematical expectation, the superscript $*$ holds for complex conjugate and r_x denotes the autocorrelation function of the stationary random process $x(n)$.

Thus the complete knowledge of the autocorrelation function of $\{v(n), n \in \mathbb{Z}\}$ is not used. Moreover, the estimation of $r_x(k)$ is done on the observed samples $x(n)$. Obviously, due to decimation, the length of the random sequence $x(n)$ is M times shorter than the original signal $v(n)$. Therefore, classical autocorrelation estimators on $x_m(n)$ result in an increase of the autocorrelation variance by a factor of M. Anyway, even if the autocorrelation estimation is done directly on $v(n)$, only the lags which are multiples of M are estimated. Therefore, the drawback on the increase of estimation variance remains.

These considerations have lead us to propose another subband spectral estimator. It is based on the idea of using the M subseries of (1) in decreasing the variance of the estimated PSD.

The principle of the proposed method is depicted in Fig. 2. Rather than modeling one of the subseries $x_m(n)$ as a linear combination of $x_m(n-1)$, $x_m(n-2), \ldots, x_m(n-p)$ as classical AR modeling does, the model studied in the present paper predicts $x_m(n)$ as a linear combination of $x_{m+d}(n)$, $x_{m+d}(n-1), \ldots, x_{m+d}(n-p+1)$. Due to its obvious link with AR modeling, we call this model a TS-AR model. More precisely, the TS-AR prediction of a subband for a given subseries m and a given time shift d can be written as

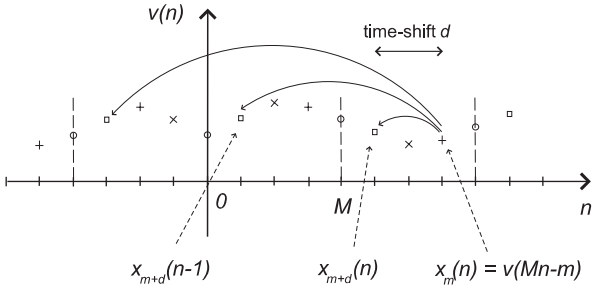

Fig. 2. Principle of the proposed method: the sample "+" is modeled as a linear combination of the past samples "□".

$$\widehat{x}_m(n) = -\sum_{k=1}^{p} a_{k,d} x_{m+d}(n-k+1) \quad \forall m \in \{0, \ldots, M-1\}, \tag{3}$$

where p is the model order. This predictor is causal provided $d \geqslant 1$. Note that the prediction coefficients do not depend on the subseries index m. This will be shown in the next section. The prediction error is defined as follows

$$e_{m,d}(n) = x_m(n) - \widehat{x}_m(n) = x_m(n) + \sum_{k=1}^{p} a_{k,d} x_{m+d}(n-k+1). \tag{4}$$

It should be noted that (3) allows to derive many models depending on the value of the time-shift d. For $d = M$, the TS-AR model is equivalent to the classical AR model. It is also interesting to highlight that the TS-AR prediction (3) leads to consider a specific prediction model of the filtered (and non-decimated) random sequence $\{v(n), n \in \mathbb{Z}\}$. Indeed, using (1), (3) can be written as

$$\widehat{v}(Mn) = -\sum_{k=1}^{p} a_{k,d} v(Mn - d - (k-1)M). \tag{5}$$

As in the case of classical AR modeling, an extra assumption must be made on the prediction error $e_{m,d}(n)$ to define a TS-AR process. Therefore, a set of decimated stationary random sequences $\{x_m(n), m \in \{0, \ldots, M-1\}, n \in \mathbb{Z}\}$ is defined as ideal TS-AR process with parameter $d > 0$ if and only if

- It exists a set of p time-independent coefficients $\{a_{k,d}, k = 1, \ldots, p\}$ such that $x_m(n)$ follows the time recursion:

$$x_m(n) = -\sum_{k=1}^{p} a_{k,d} x_{m+d}(n-k+1) + e_{m,d}(n). \tag{6}$$

- The error $e_{m,d}(n)$ is decorrelated with all past signal samples $x_{m+d}(n-k+1)$ for all $k > 0$, i.e.

$$r_{e_{m,d} x_{m+d}}(s) = 0 \quad \forall s > 0. \tag{7}$$

Note that in the classical case of an AR process, this definition is quite the same [13], except that the assumption (ii) states that the AR model error is an ideal white noise. However, in the case of an AR process, the whiteness of the error can be shown to be equivalent to the decorrelation between the error and the past AR samples. The TS-AR definition involves only a decorrelation assumption since this

property is necessary to derive parameter estimation as it will be shown in the next section. Moreover, it will be shown that, even with this decorrelation assumption, the TS-AR error is not white as in the classical AR case.

In order to illustrate the interest of this model, it can be remarked that typical examples of signals following (3) with null error (special case of TS-AR processes) are pure sinusoids. Indeed, if $v(n)$ is a sinusoidal signal, the set of decimated signals $x_m(n)$ matches the following equations:

$$x_m(n) = B\sin(2\pi v_0(Mn - m) + \Phi) \quad \forall m \in \{0, \ldots, M-1\}, \tag{8}$$

where B is the amplitude, v_0 is the normalized frequency and Φ is the corresponding phase, uniformly distributed between 0 and 2π for instance. Then, if v_0 is not multiple of π/M, $x_m(n)$ is a true TSAR signal of order 2 with null error for all parameters d because $\forall d$ and $\forall m \in \{0, \ldots, M-1\}$:

$$x_m(n) = -\frac{\sin(2\pi v_0(M+d))}{\sin(2\pi v_0 M)} x_{m+d}(n) + \frac{\sin(2\pi v_0 d)}{\sin(2\pi v_0 M)} x_{m+d}(n-1). \tag{9}$$

4. Model parameter estimation

4.1. Pseudo Yule–Walker equations

In order to estimate the prediction coefficients of (3), a natural criterion is the least square error minimization. Thus, minimizing the model error power (4) yields the following p equations:

$$\forall s \in \{1, \ldots, p\}, \quad E[e_{m,d}(n) x^*_{m+d}(n+1-s)] = 0. \tag{10}$$

Using (4) and (2), one obtains

$$\forall s \in \{1, \ldots, p\}, \quad r_v(Ms + d - M) + \sum_{k=1}^{p} a_{k,d} r_v(M(s-k)) = 0. \tag{11}$$

These equations correspond to the p first pseudo Yule–Walker equations for TS-AR signals. The use of (7) allows to demonstrate it for all $s>0$: multiplying each side of (4) by $x^*_{m+d}(n-s+1) = v^*(M(n-s+1) - m - d)$ and taking the mathematical expectation yields

$$r_v(Ms + d - M) + \sum_{k=1}^{p} a_{k,d} r_v(M(s-k)) = E[e_{m,d}(n) x^*_{m+d}(n-s+1)] \quad \forall s > 0. \tag{12}$$

Either for a true TS-AR signal or for an unspecified random sequence modeled as a TS-AR signal, the error is orthogonal to past signal samples (7) thus

$$r_v(Ms + d - M) + \sum_{k=1}^{p} a_{k,d} r_v(M(s-k)) = 0 \quad \forall s > 0. \tag{13}$$

For a given time-shift $d \geqslant 1$, the set of equations (11) can be written in a matrix form. In order to estimate the prediction coefficients $\{a_{k,d}, k = 1, \ldots, p\}$, the autocorrelation function $r_v(k)$ is first estimated and the following equation is solved:

$$\widehat{R}.\underline{\widehat{a}}_d = -\underline{\widehat{r}}_d \tag{14}$$

with

$$\widehat{R} = \begin{pmatrix} \widehat{r}_v(0) & \widehat{r}_v(-M) & \ldots & \widehat{r}_v(-M(p-1)) \\ \widehat{r}_v(M) & \widehat{r}_v(0) & & \\ \vdots & & \ddots & \\ \widehat{r}_v(M(p-1)) & & & \widehat{r}_v(0) \end{pmatrix}, \tag{15}$$

where $\underline{\widehat{r}}_d = [\widehat{r}_v(d)\ \widehat{r}_v(M+d)\ \ldots\ \widehat{r}_v(M(p-1)+d)]^{\mathrm{T}}$, $\underline{\widehat{a}}_d = [\widehat{a}_{1,d}\ \widehat{a}_{2,d}\ \ldots\ \widehat{a}_{p,d}]^{\mathrm{T}}$, $\widehat{r}_v(k)$ and $\widehat{a}_{k,d}$ denote respective estimates of $r_v(k)$ and $a_{k,d}$.

Note that $\widehat{R}$ and $\underline{\widehat{r}}_d$ are independent of the subseries index m, which ensures that the $a_{k,d}$ coefficients and their estimates $\widehat{a}_{k,d}$ are also independent of m, as it was set in the model (3). This is also shown by the equivalent model in (5).

Moreover, the correlation matrix $\widehat{R}$ has a Toeplitz structure, which allows the use, for example, of the Levinson–Durbin algorithm for a fast inversion of the system (see [13, p. 213]). Note also that this matrix does not depend on d. Actually, $\widehat{R}$ is the classical autocorrelation matrix of AR estimation from $x_m(n)$. This yields two advantages. First, the computation of its inverse will be done only once, even if several time shifts are used. Second, the solving of (14) can be achieved using any AR algorithm. Thus, several TS-AR models can be defined, each characterized by a shift $d \geqslant 1$.

In order to build a spectral analysis tool, statistical properties of the prediction error $e_{m,d}(n)$ are studied in the next section.

4.2. Prediction error statistical properties

The mean of the linear prediction error can be derived using (4). If the input random process $\mathscr{U} = \{u(n), n \in \mathbb{Z}\}$ is assumed to be zero-mean, obviously, the linear prediction error of the TS-AR model is zero-mean. For $s \geqslant 0$, using (4), its autocorrelation function is of the form

$$\begin{aligned} r_{e_{m,d}}(s) &= E[e_{m,d}(n)e^*_{m,d}(n-s)] \\ &= r_v(Ms) + \sum_{k=1}^{p} a_{k,d} r_v(M(s-k+1)-d) \\ &\quad + \sum_{q=1}^{p} a^*_{q,d} r_v(M(s+q-1)+d) \\ &\quad + \sum_{q=1}^{p} a^*_{q,d} \sum_{k=1}^{p} a_{k,d} r_v(M(s+q-k)). \end{aligned} \tag{16}$$

Eq. (13) yields

$$\begin{aligned} &\forall s \geqslant 0, \quad \sum_{k=1}^{p} a_{k,d} r_v(M(s+q-k)) \\ &\quad = -r_v(M(s+q-1)+d). \end{aligned} \tag{17}$$

Finally, (16) can be written in a simplified manner, $\forall s \geqslant 0$,

$$r_{e_{m,d}}(s) = r_v(Ms) + \sum_{k=1}^{p} a_{k,d} r_v(M(s-k+1)-d). \tag{18}$$

Eq. (18) shows that this error autocorrelation function does not depend on the subseries index m. Therefore, in what follows, it will be denoted by $r_{e_d}(s)$. By replacing the theoretical autocorrelation samples $r_v(k)$ by their estimates $\widehat{r}_v(k)$, we obtain the expression for estimation of $r_{e_d}(s)$. Contrarily to the case of classical AR modeling, the prediction error of TS-AR modeling is no longer white, even if the modeled signal is a pure TS-AR signal matching (7) (except in the case $d = M$ which is equivalent to AR modeling). In all other cases, for $d \in \{1, 2, \ldots, M-1\}$ and $d \geqslant M+1$, the autocorrelation $r_{e_d}(s)$ can be estimated until at least lag p using (18). This allows to estimate the PSD of any subseries $x_m(n)$ (all subseries have the same PSD), as it is shown in Section 5.

4.3. Autocorrelation estimation

The TS-AR coefficient estimation, like in classical AR modeling, is based on the estimates of the autocorrelation samples of $r_x(k)$. Within the frame of subband processing, several estimators may be used. Considering one branch of the filterbank depicted in Fig. 1, there are a priori at least three different ways to estimate the autocorrelation function $r_x(k)$.

A first estimator can be derived, based on a two-step procedure. First, $\widehat{r}_u(k)$ is estimated from the N available samples, using a traditional autocorrelation estimator (biased or unbiased). Second, considering that the subband filters are of finite impulse response (FIR) type, the autocorrelation of the filtered signal $v(n)$ is of the form:

$$r_v(k) = \sum_{s=-(L-1)}^{L-1} r_h(s) r_u(k-s),$$

with

$$r_h(s) = \sum_{n=0}^{L-1} h(n)h^*(n-s), \tag{19}$$

$h(n)$, $n = 0, \ldots, L-1$ being the impulse response of the considered filter. Note that the non-overlapping procedure proposed in [11,12] induces FIR filters. This leads to the first estimator:

$$\overset{\vee}{r}_x(k) = \sum_{s=-(L-1)}^{L-1} r_h(s) \widehat{r}_u(Mk-s). \tag{20}$$

A second estimator can be proposed, based on the classical estimation of $\widehat{r}_v(k)$ from $v(n)$ samples:

$$\breve{r}_x(k) = \widehat{r}_v(Mk). \tag{21}$$

A third estimator can also be considered. It consists of estimating M autocorrelation functions, $\widehat{r}_{x_m}(k)$, from the M available subseries $x_m(n)$, using a classical estimator and then averaging them out, that is,

$$\widetilde{r}_x(k) = \frac{1}{M} \sum_{m=0}^{M-1} \widehat{r}_{x_m}(k). \tag{22}$$

Obviously, all three estimators have the same mean. To compare them, a variance study is done in Appendix A which shows that in the case of Gaussian processes

$$Var[\widetilde{r}_x(k)] = Var[\breve{r}_x(k)]. \tag{23}$$

However, no interesting expression of the variance of $\overset{\vee}{r}_x(k)$ can be found in order to compare it with the variance of the other two estimators. Therefore, the choice between these three estimators is based on the aliasing-free condition in the center of the subbands, which is strictly satisfied by the first estimator (see [12]) and only approximately by the other two estimators. For this reason, the first estimator $\overset{\vee}{r}_x(k)$ has been used in this work.

5. Spectral analysis based on TS-AR modeling

In this section, the analysis filterbank of Fig. 1 is assumed to be ideal, i.e. composed of infinitely sharp bandpass filters. For $d \geqslant 1$, the TS-AR model is given by $\forall n \in \mathbb{Z}$, $\forall m \in \{0, \ldots, M-1\}$,

$$x_m(n) = -\sum_{k=1}^{p} a_{k,d}\, x_{m+d}(n-k+1) + e_{m,d}(n). \qquad (24)$$

To remain coherent with (1), the index $m+d$ should be less than M. Note that if m and d are such that $m+d \geqslant M$, there exists a unique couple $(q,r) \in \mathbb{N}^2$ such that $m+d = Mq+r$ with $0 \leqslant r \leqslant M-1$. Therefore, in this case, (24) can be modified using

$$x_{m+d}(n) = x_r(n-q). \qquad (25)$$

(This comes directly from (1).) Then (24) leads to the PSD of the TS-AR model, denoted by $S_{\text{TS-AR}}(f)$:

$$S_{\text{TS-AR}}(f) = \frac{S_{e_d}(f)}{\left|1 + \sum_{k=1}^{p} a_{k,d} \mathrm{e}^{-\mathrm{i}2\pi f(k+(d-1)/M)}\right|^2}, \qquad (26)$$

where $S_{e_d}(f)$ denotes the PSD of the model error. This PSD $S_{e_d}(f)$ can be estimated using (18). Therefore, TS-AR modeling can be applied on each branch of the analysis filterbank of Fig. 1. Each of these TS-AR models leads to an estimation of the PSD of $\{u(n), n \in \mathbb{Z}\}$ within the considered subband. A specific processing, such as the one proposed in [11,12], has to be applied in order to avoid any overlapping between the different subbands. Then, the whole PSD $S_u(f)$ can be estimated subband after subband, from the estimates of $S_{\text{TS-AR}}(f)$ in each subband. Obviously, the procedure can be implemented in parallel, allowing the PSD estimation on the M subbands simultaneously.

6. Main benefits of the method

It is well-known that the variance of the autocorrelation estimators increases when the lag tends to the available number of samples N. Since $r_x(k) = r_v(Mk)$, there are several ways to interpret the second drawback (D_2) of subband parametric spectral estimation when compared to the fullband estimation.

Let p_{full} and p denote the model orders chosen for the fullband process and for the subband processes, respectively. Note that in AR-based spectral estimation methods, these generally correspond to the number of autocorrelation samples to be estimated (except for least-square based algorithms). Then, if $p = p_{\text{full}}$, there is an increase of the global variance of the estimated autocorrelation vector,

$$\| Var[\widehat{\underline{r}}_{\text{sub}}] \| > \| Var[\widehat{\underline{r}}_{\text{full}}] \|$$

where $\widehat{\underline{r}}_{\text{sub}}$ and $\widehat{\underline{r}}_{\text{full}}$ represent the respective estimated autocorrelation vectors of the subband and fullband processes. However, if $p = p_{\text{full}}/M$, the estimation variance will be the same. But, in this case, there is information reduction due to the fact that the vector $\underline{r}_{\text{sub}}$ has a size M times smaller than the vector $\underline{r}_{\text{full}}$. Eq. (11) shows that TS-AR method allows the use of autocorrelation samples of signal $v(n)$ with indices $Mk+m$, $m \neq 0$, contrarily to classical subband AR estimation. Note that these $v(n)$ autocorrelation samples (for $m \in \{1, \ldots, M-1\}$) correspond to the intercorrelation between two decimated subseries in the same subband:

$$\begin{aligned} r_{x_{m_1} x_{m_2}}(k) &= E[x_{m_1}(n) x^*_{m_2}(n-k)] \\ &= r_v(Mk + m_2 - m_1). \end{aligned} \qquad (27)$$

Therefore, the TS-AR spectral estimation may help in reducing drawback (D_2) by finding a way to use this additional information.

Moreover, as the number of available autocorrelation samples is often limited by the number of signal samples N, it seems better to use an order $p = p_{\text{full}}/M$ for a fair comparison of fullband and subband parametric spectral estimation. This is the reason of our choice in the present paper.

Classical parametric spectral estimation from subbands has already been shown to be beneficial for several reasons, as stated by properties B_1–B_4 (see Section 1) and these benefits are preserved with this new time-shift procedure.

Considering property B_1, TS-AR modeling leads to a decrease of the condition number since the matrix $\widehat{R}$ is the same as in the case of classical AR modeling in subbands (see Section 4).

The B_2 property is still verified with TS-AR modeling because frequency spacing and local

signal to noise ratio (SNR) increases are only due to the decimation and filtering operations.

Property B_3 concerns PSD of signals and not PSD of the modeling error. When a given signal $y(n)$ satisfies the Paley–Wiener condition ($\int_{-1/2}^{1/2} |\log S_y(f)|\,\mathrm{d}f < \infty$: see [14]), a realistic spectral flatness measure can be used as in [15]

$$\mathscr{F}_y = \frac{\exp[\int_{-1/2}^{1/2} \ln S_y(f)\,\mathrm{d}f]}{\sigma_y^2}, \tag{28}$$

where σ_y^2 is the signal variance. When using an ideal filterbank (uniform or not), it can be shown that subband signals have a flatter PSD than the original one in geometric mean [10]. As this property does not depend on the kind of modeling, it remains valid with TS-AR modeling.

Property B_4 has been theoretically shown by Rao and Pearlman in [7]. More precisely, the analyzed fullband signal $u(n)$ is assumed to be a real AR(q) process decomposed with an ideal filterbank (infinitely sharp bandpass filters). Let ρ_u denote the linear prediction error power associated to the fullband classical AR modeling and ρ_j the one associated to each of the M subband AR modelings. Then the following inequality is verified:

$$\rho_u \geqslant \sum_{j=0}^{M-1} \rho_j \quad \text{for } p \leqslant q \tag{29}$$

with equality if and only if $u(n)$ is a white noise. In the asymptotic case ($M \to +\infty$), this means that the linear prediction error powers ρ_j must tend to 0 whatever j for the series to stay convergent. This result shows the advantage of subband decomposition for AR modeling when a linear prediction error power criterion is considered. The result has been obtained for an ideal filterbank and for classical AR modeling. Since the TS-AR model is equivalent to the classical AR modeling for $d = M$, this property will be investigated in the next section. The next section presents simulation results that compare the TS-AR spectral estimation with that of classical fullband and subband AR spectral estimation.

7. Simulation results

7.1. Simulation context and time-shift d selection

TS-AR modeling has been presented without any filter implementation consideration. However, it is well-known that non-ideal filters result in spectral overlapping. In order to avoid this problem, a specific procedure has been proposed in [12]. The solution is composed of a warping device, placed before the filterbank which is chosen as a bank of FIR modulated comb filters of order L. The squared modulus of the transfer function of the jth filter is given by

$$|\mathscr{H}_j(\mathrm{e}^{\mathrm{i}2\pi f})|^2 = \begin{cases} \dfrac{1}{L}\dfrac{\sin^2 \pi(f-F_j)L}{\sin^2 \pi(f-F_j)} & \text{if } f \neq F_j, \\ L & \text{if } f = F_j, \end{cases} \tag{30}$$

where j is the subscript of the considered subband ($j \in \{0, \ldots, M-1\}$) and $F_j = j(0.5/M) + 0.25/M$ is the center of the jth subband B_j. Thus, the transfer function is periodically equal to zero around F_j with period $1/L$. This is the alias-free condition for spectral estimation from subbands when using the procedure described in [12]. For sake of clarity, a short description of the method is given hereafter.

Let $v \in [0, 1]$ be the frequency to which the PSD $S_u(v)$ is to be estimated. In Fig. 1, the filterbank is preceded by a warping device (complex exponential $\mathrm{e}^{\mathrm{i}2\pi\Delta f n}$ with $\Delta f = F_j - v$). The choice of Δf allows to warp the spectral estimation at frequency v to the center of its corresponding subband F_j, where aliasing is canceled by the properties of the analysis filterbank. Thus, the algorithm used to lead simulations presented in this paper is summarized below:

For $d \geqslant 1$ and for each frequency v:

1. Subband selection $j = \mathrm{Int}[2Mv]$ (with Int[.] the integer part operator).
2. Autocorrelation estimation $\widehat{r}_x(k)$ using the first estimator presented in Section 4.3.
3. TS-AR parameter estimation using classical AR algorithm to solve (14).
4. Error spectral estimation using (18) to derive the error spectrum at MF_j: $S_{e_d}(MF_j)$.
5. TS-AR spectral estimation using (26) to derive $S_{\text{TS-AR}}(MF_j)$.
6. Reconstruction of the spectral estimate of the input signal:

   $\widehat{S}_u(v) = MS_{\text{TS-AR}}(MF_j)$.

   This last result comes from the particular properties of the comb filterbank and is valid for filter order L multiple of M (see [12]).

While varying the above frequency v between 0 and 1 (or 0.5 for real signals) with a given frequency

step, the whole PSD of the original signal $u(n)$ can then be reconstructed.

The main parameters of this TS-AR algorithm are the number of subbands M, the subband modeling order p and the time-shift d. In practice, the selection of M and p can be done in the same way than classical subband AR modeling:

1. Fullband order selection: find the best order p_{full} (using classical criterions as MDL or AIC for example) for fullband AR modeling. As explained previously, M and p are then linked by the following relationship:

   $Mp = p_{\text{full}}$.
2. Find the larger decimation ratio M for which the subband order $p = p_{\text{full}}/M$ is sufficient for modeling the subband signal well (in the sense of the same criterion as above).

The choice of the time-shift d if of great importance because, from the simulations below, it is clear that it affects the performance. It is difficult to choose because its optimal value (under a given criterion) depends on the spectrum of the analyzed signal $u(n)$. The first thing to consider for the choice of this parameter is that d must not be too large for signals with continuous spectrum. Indeed, in this case, the autocorrelation $r_v(k)$ tends to 0 when $k \to +\infty$. As a consequence, when $d \to +\infty$, $\underline{r}_d = [r_v(d)\ r_v(M+d) \ldots r_v(M(p-1)+d)]^{\mathrm{T}}$ tends to null vector. From (14), the following result is obtained:

$$\underline{a}_d \to \underline{0} \quad \text{when } d \to +\infty. \tag{31}$$

In practice, when $d \gg M$, the TS-AR coefficients vanish and the spectrum of the error $S_{e_d}(f)$ tends to the signal spectrum: the TS-AR modeling becomes inefficient.

The proposed solution is to derive the TS-AR estimate for $d = 1, 2, \ldots, M+1$ and then to choose d so that the variance of the estimator is minimum (requires several realizations). Also, in the case of signals with line spectrum, it is not useful to choose a large time-shift d because the TS-AR estimate will have a periodic behavior, depending on the frequencies.

All simulations presented in what follows are driven within this context. The next section presents the performance of the TS-AR method on two kinds of signals, with continuous and with line spectra. Comparisons between fullband, classical subband and TS-AR spectral analysis are given. In particular, the LPE behavior is studied for the three kinds of spectral analysis. Dealing with this, a remark has to be done. For fullband spectral estimation, there is only one autocorrelation matrix whatever the frequency of interest and therefore, the LPE power is not frequency dependant. However, subband and TS-AR spectral estimations are made using the non-overlapping procedure of [12], as detailed above. This results in a specific spectral estimation for each frequency of interest. This is the reason why the LPE power is frequency dependent in the case of TS-AR and classical subband spectral estimation.

7.2. Signals with continuous spectrum

Simulations are done in the case of a pure MA signal with continuous spectrum. For each realization, the simulated signal is obtained by finite impulse response (FIR) filtering of a white noise

$$\mathscr{U}(z) = \mathscr{B}(z)\mathscr{W}(z), \tag{32}$$

where $w(n)$ is a white noise with power $\sigma_w^2 = 1.2$ and the MA coefficients are given by

$$[1.0000, -0.1837, 0.5373, \\ -0.3252, 0.4351, 0.1419, 0.0174].$$

These coefficients are chosen so that the theoretical spectrum is neither too peaky nor too flat. The chosen model order for fullband spectral estimation is $p_{\text{full}} = 16$ and $p = 4$ for subband spectral estimation (for AR and TS-AR modeling). The filterbank is uniform with $M = 4$ subbands and a decimation ratio of M.

The simulations presented below are obtained using 100 Monte-Carlo runs with the above parameters. In order to compare the performance of TS-AR as well as the standard subband and fullband AR spectral estimations, linear prediction error (LPE) spectra are plotted in Fig. 3. In the case of TS-AR modeling, the error spectrum for the different considered time-shifts d, $S_{e_d}(MF_j)$, is displayed rather than the LPE power since it is a better measure of the performance (see (26)). In Table 1, these LPEs are averaged on each subband and averaged spectrum variances are also included to demonstrate how the chosen delay affects final performances.

As can be seen in Table 1 and Fig. 3, the TS-AR method provides better results in terms of LPE power than the fullband and standard subband spectral estimation methods for short time-shifts ($d < M$). The reason is the following: in Fig. 2, it can

be seen that d represents physically the prediction lag: equidistant samples $x_{m+d}(n-k)$ for different k are used to predict sample $x_m(n)$ which is separated by d samples from the others. As it is more difficult to predict something far away, TS-AR provides lower LPE than classical subband method ($d = M = 4$) for $d < M$. On the other hand, Table 1 and Fig. 3 also show that the parameter giving minimum variance for the spectral estimator (for the considered MA signal) is $d = 5$ and not 1 (which gives minimum LPE power). When dealing with continuous spectrum signals, it is often desirable to get low variance estimates and this makes variance a good performance criterion. This raises an apparent inconsistency between the optimization criterion and the final performance measure which comes from the fact that these two criteria are equivalent in the case of classical AR modeling (fullband or subband) but not in the case of TS-AR modeling. Rather than the LPE power, a better optimization criterion for TS-AR parameter estimation would be the whiteness of the modeling error. However, this would result in a high computational cost of the algorithm. Using LPE power minimization allows to use classical algorithms to solve (14) and reduces the cost of the overall algorithm which is anyway complex (for each frequency ν, (14) has to be solved).

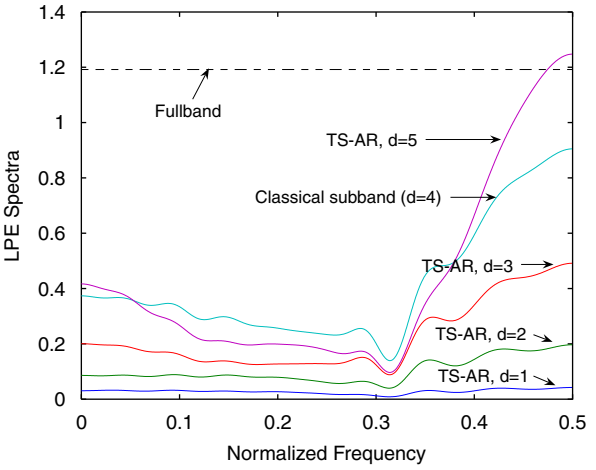

Fig. 3. Linear prediction error spectra.

In the following, parameter $d = 5$ is chosen to illustrate TS-AR performances versus classical fullband and subband ones since it minimizes the estimation variance. The bias and mean square errors (MSE) of the spectral estimates are displayed in the next figures. The aim is to compare the three different spectral estimators: the classical fullband and subband AR estimations and the TS-AR estimator for $d = 5$. More precisely, knowing the true spectrum $S(f)$ and naming $\widehat{S}_i(f)$, $i = 1, \ldots, 100$ the different realizations of a given spectral estimator, the means and variances are obtained as follows:

$$\overline{\widehat{S}}(f) = \frac{1}{100}\sum_{i=1}^{100}\widehat{S}_i(f),$$

$$Var[\widehat{S}(f)] = \frac{1}{100}\sum_{i=1}^{100}(\widehat{S}_i(f) - \overline{\widehat{S}}(f))^2. \tag{33}$$

Table 1
Linear prediction error powers (continuous spectrum)

		Subband $j = 0$	Subband $j = 1$	Subband $j = 2$	Subband $j = 3$	Means on all Subbands
Fullband	σ^2			1.1918		
	$Var[S]$			0.3431		
TS-AR, $d = 1$	σ^2	0.0306	0.0259	0.0189	0.0353	0.0277
	$Var[S]$	0.3765	0.2087	0.0987	0.9489	0.4082
TS-AR, $d = 2$	σ^2	0.0848	0.0778	0.0805	0.1679	0.1027
	$Var[S]$	0.3396	0.1874	0.0885	0.9118	0.3818
TS-AR, $d = 3$	σ^2	0.1766	0.1299	0.1759	0.4106	0.2233
	$Var[S]$	0.3338	0.1789	0.0855	0.8925	0.3727
TS-AR, $d = 4$	σ^2	0.3462	0.2649	0.2884	0.7413	0.4102
(Classical subband)	$Var[S]$	0.3235	0.1699	0.0824	0.8726	0.3621
TS-AR, $d = 5$	σ^2	0.3301	0.1986	0.2185	0.9364	0.4209
	$Var[S]$	0.3282	0.1704	0.0819	0.8668	0.3618

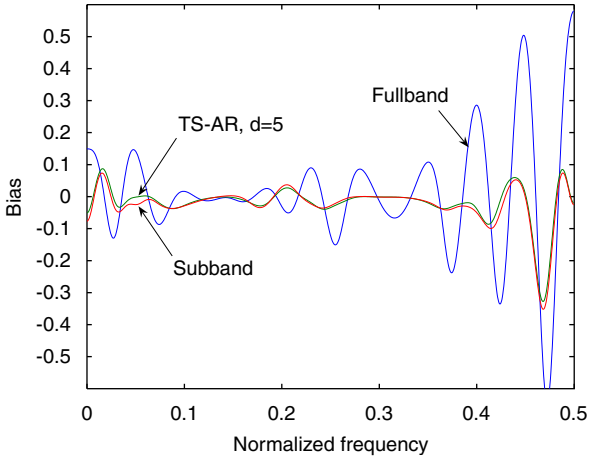

Fig. 4. Spectrum biases versus frequency.

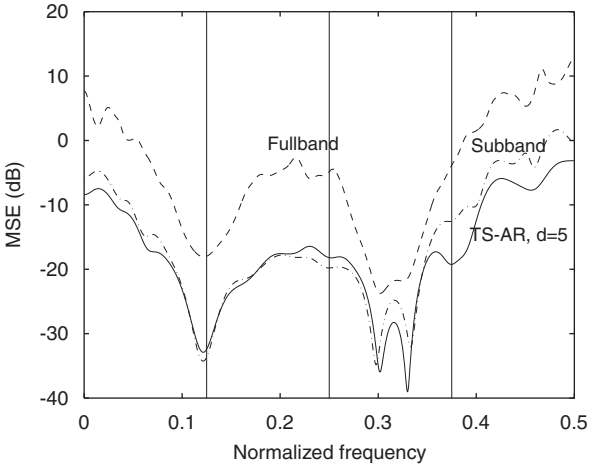

Fig. 5. Spectrum MSE versus frequency.

Using 100 Monte-Carlo runs, the bias $S(f) - \overline{\widehat{S}}(f)$ and the MSE $\mathrm{MSE}(f) = (S(f) - \overline{\widehat{S}}(f))^2 + Var[\widehat{S}(f)]$ are plotted in Figs. 4 and 5.

Even if the TS-AR and classical subband biases are quasi-superimposed, the variance is smaller in the case of TS-AR modeling than in the case of fullband or subband AR modeling. These results show the advantage of the TS-AR method in reducing the estimation variance of continuous spectra, even if this improvement is weak.

7.3. *Signals with line spectrum*

In this section, the analyzed signal is assumed to be a pure sinusoid embedded in white noise with a random phase ϕ uniformly distributed between 0 and 2π.

$$u(n) = A\sin(2\pi f_0 n + \phi) + b(n), \tag{34}$$

where $A = 1$, $f_0 = 0.1$ (normalized frequency) and $b(n)$ is a white noise with power $\sigma_b^2 = 0.05$ (SNR $= 10$ dB). In the case of line spectrum signals, the final performance measure is the variance of the frequency estimate. Therefore, a specific frequency estimator based on TS-AR modeling, is proposed for this kind of signals.

The idea is to use all the TS-AR spectral estimators and compute an averaged TS-AR frequency estimate. More precisely, the frequency estimate $\widehat{f}_{0d}$ of the sinusoid frequency f_0 is determined on each separate spectral estimate for $d = 1, \ldots, 5$, given by the maximum of the estimated spectrum or more simply, zeroing the denominator of (26). Then these estimations $\widehat{f}_{0d}$, $d = 1, \ldots, 5$ are averaged to form the final TS-AR estimate $\widehat{f}_0$. Indeed, in the case of line-spectrum signals, the frequency estimate is the one zeroing the denominator of (26) and is therefore independent of extra errors coming from the estimation of the error spectrum $S_{e_d}(f)$. This is not the case when estimating the spectrum of a MA signal using (26) as in the previous subsection.

The bias and variance of $\widehat{f}_0$ are plotted versus the realization number in Figs. 6 and 7 using 500 Monte-Carlo runs. The goal is still to compare the three methods (classical AR fullband and subband and TS-AR). As in the case of continuous spectra, the variance and bias of the frequency estimates are smaller with TS-AR modeling.

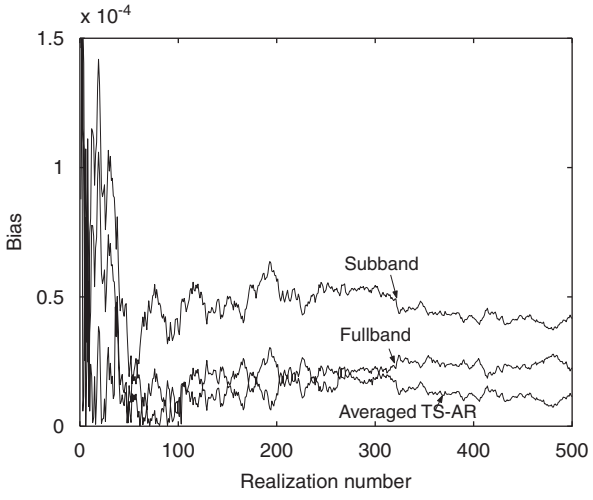

Fig. 6. Bias of $\widehat{f}_0$ versus the realization number.

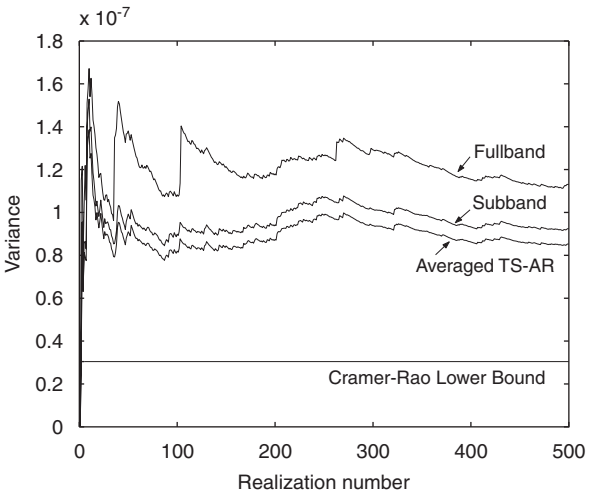

Fig. 7. Variance of $\widehat{f_0}$ versus the realization number.

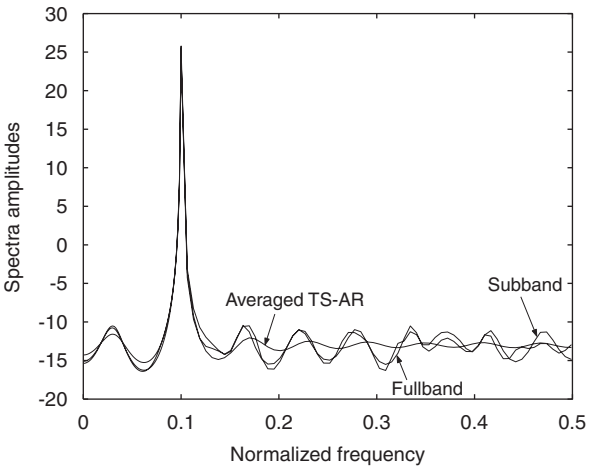

Fig. 8. Estimated spectra averaged on 500 Monte-Carlo runs.

Moreover, this idea of using all TS-AR models with different values of d can be extended to propose a spectral estimator. Rather than choosing one value of d, which is a non-trivial task as we have shown in the previous section, we propose to form an averaged TS-AR spectral estimator defined as the average of all TS-AR spectral estimators (26) for different values of d. The improvement brought by TS-AR modeling is shown in Fig. 8 where fullband AR, subband AR and averaged TS-AR spectra are displayed, showing that the TS-AR spectral estimator has much smaller variance.

8. Conclusion

The aim of this paper was to propose a spectral estimation improvement within the field of subband estimation. A new AR-based model was proposed as a way to exploit the a priori information provided by the knowledge of the autocorrelation samples of the filtered signal $v(n)$ on one branch of the filterbank, or, which is equivalent, by the knowledge of the intercorrelation between the different subseries $x_m(n)$ obtained after the M-fold decimator.

After the definition of this new modeling, TS-AR parameter estimation has been presented, using a classical minimum error power criterion. It should be noted that this criterion is not optimal in the sense of the error whiteness as for classical AR modeling. However, this criterion allows the use of classical AR parameter estimation algorithms for TS-AR parameter estimation.

Moreover, different ways to estimate the autocorrelation function of the decimated sub-signals have been presented and studied, showing that two out of them are equivalent at first and second order. Simulations were performed both in the case of signals with continuous spectrum and in the case of signals with line spectrum. They compared the TS-AR spectral estimation and the classical subband and fullband AR methods. In the case of signals with continuous spectra, the improvement brought by TS-AR modeling is weak. Moreover, selection of the time-shift d is difficult in this case as it requires several realizations of the studied signal. The real interest of TS-AR modeling appears for frequency estimation (case of line spectra): the use of an averaging of the TS-AR estimates with different parameter d values reduces significantly the bias and variance of the frequency estimate and no time-shift selection is necessary.

Acknowledgments

The authors would like to thank Petar M. Djurić from Stony Brook University, New York (Department of Electrical and Computer Engineering) for the help provided through many discussions about the paper.

Appendix A. Variance of the autocorrelation estimators

The aim of this appendix is to study the autocorrelation estimators $\breve{r}_x(k)$ and $\widetilde{r}_x(k)$ defined, respectively, in (21) and (22). Obviously, both estimators have the same mean.

The variance of the first estimator is

$$Var[\breve{r}_x(k)] = Var[\widehat{r}_v(Mk)]. \tag{A.1}$$

Assuming the process to be Gaussian allows to derive simplified expression of the variance of classical autocorrelation estimators [16]:

$$\begin{aligned} Var[\widehat{r}_{x_m}(k)] &= C \sum_{l=-\infty}^{+\infty} (r_{x_m}^2(l) + r_{x_m}(l+k) r_{x_m}(l-k)) \\ &= C \sum_{l=-\infty}^{+\infty} (r_v^2(Ml) \\ &\quad + r_v(Ml+Mk) r_v(Ml-Mk)), \end{aligned} \tag{A.2}$$

where $C = 1/N_s$ for the biased estimator and $C = N_s/(N_s-k)^2$ for the unbiased estimator, N_s being the number of available samples for signal $x_m(n)$. Let $J_M(l)$ denote the following function:

$$J_M(l) = \begin{cases} 1 & \text{if } l \text{ is multiple of } M, \\ 0 & \text{else.} \end{cases} \tag{A.3}$$

As $J_M(l)$ is periodic, another expression can be derived by writing its Fourier series expansion

$$J_M(l) = \frac{1}{M} \sum_{s=0}^{M-1} e^{i(2\pi ls/M)}. \tag{A.4}$$

Eq. (A.2) can be rewritten as

$$\begin{aligned} &Var[\widehat{r}_{x_m}(k)] \\ &\quad = C \sum_{l=-\infty}^{+\infty} (r_v^2(l) + r_v(l+Mk) r_v(l-Mk)) J_M(l) \\ &\quad = \sum_{s=0}^{M-1} \frac{C}{M} \sum_{l=-\infty}^{+\infty} [(r_v(l) e^{i(\pi ls/M)})^2 \\ &\qquad + (r_v(l+Mk) e^{i(\pi(l+Mk)s/M)}) \\ &\qquad \times (r_v(l-Mk) e^{i(\pi(l-Mk)s/M)})]. \end{aligned} \tag{A.5}$$

Introducing $w_s(n) = v(n) e^{i(\pi ns/M)}$, yields

$$\begin{aligned} Var[\widehat{r}_{x_m}(k)] = \sum_{s=0}^{M-1} \frac{C}{M} \sum_{l=-\infty}^{+\infty} [r_{w_s}^2(l) \\ + r_{w_s}(l+Mk) r_{w_s}(l-Mk)]. \end{aligned} \tag{A.6}$$

Signal $w_s(n)$, constructed from non-decimated signal $v(n)$ has MN_s samples, so that in both cases:

$$\begin{aligned} &Var[\widehat{r}_{w_s}(Mk)] \\ &\quad = \frac{C}{M} \sum_{l=-\infty}^{+\infty} [r_{w_s}^2(l) + r_{w_s}(l+Mk) r_{w_s}(l-Mk)]. \end{aligned}$$

Hence

$$Var[\widehat{r}_{x_m}(k)] = \sum_{s=0}^{M-1} Var[\widehat{r}_{w_s}(Mk)]. \tag{A.7}$$

Moreover, $Var[\widehat{r}_{w_s}(k)] = Var[\widehat{r}_v(k)]$, $\forall k$ independently of s, so

$$Var[\widehat{r}_{x_m}(k)] = M Var[\widehat{r}_v(Mk)] \tag{A.8}$$

and

$$Var[\widetilde{r}_x(k)] = Var[\widehat{r}_v(Mk)]. \tag{A.9}$$

Finally, using (A.1) and (A.9) the variances of both estimators $\breve{r}_x(k)$ and $\widetilde{r}_x(k)$ are shown to be identical:

$$Var[\widetilde{r}_x(k)] = Var[\breve{r}_x(k)]. \tag{A.10}$$

References

[1] A. Charbonnier, J.P. Petit, Sub-band adpcm coding for high quality audio signals, in: Proceedings of IEEE ICASSP-83, vol. 5, New York, NY, USA, 1988, pp. 2540–2543.

[2] Z. Cvetkovic, J.D. Johnston, Nonuniform oversampled filter banks for audio signal processing, IEEE Trans. Speech, Audio Process. 11 (2003) 393–399.

[3] F. Baumgarte, Improved audio coding using a psychoacoustic model based on a cochlear filter bank, IEEE Trans. Speech Audio Process. 10 (7) (2002) 495–503.

[4] R. Yu, C. Ko, A warped linear-prediction-based subband audio coding algorithm, IEEE Trans. Speech Audio Process. 10 (1) (2002) 1–8.

[5] P. Broersen, S. de Waele, Time series analysis in a frequency subband, IEEE Trans. Instrumentation Meas. 52 (4) (2003) 1054–1060.

[6] O. Shentov, S. Mitra, A simple method for power spectral estimation using subband decomposition, Int. Conf. Acoustics, Speech, Signal Process. 5 (1991) 3153–3156.

[7] S. Rao, W.A. Pearlman, Analysis of linear prediction, coding, and spectral estimation from subbands, IEEE Trans. Inform. Theory 42 (4) (1996) 1160–1178.

[8] A. Tkacenko, P.P. Vaidyanathan, Sinusoidal frequency estimation using filter banks, in: Proceedings of IEEE ICASSP-01, Salt Lake City, Utah, 2001, pp. 825–828.

[9] G.H. Golub, C.F. Van-Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, MD, 1989.
[10] A. Tkacenko, P.P. Vaidyanathan, The role of filter banks in sinusoidal frequency estimation, J. Franklin Inst. 338 (5) (2001) 517–547.
[11] D. Bonacci, P. Michel, C. Mailhes, Subband decomposition and frequency warping for spectral estimation, in: Proceedings of IEEE EUSIPCO-2002, Toulouse, France, 2002, pp. 147–150.
[12] D. Bonacci, C. Mailhes, P.M. Djurić, Improving frequency resolution for correlation-based spectral estimation methods using subband decomposition, in: Proceedings of IEEE ICASSP-03, vol. 6, 2003, pp. 329–332.
[13] S.M. Kay, Modern Spectral Estimation: Theory and Applications, Prentice-Hall, Englewood Cliffs, NJ, 1988.
[14] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1991.
[15] N.S. Jayant, P. Noll, Digital Coding of Waveforms, Prentice-Hall, Englewood Cliffs, NJ, 1984.
[16] S.L.J. Marple, Digital Spectral Analysis With Applications, Prentice-Hall, Englewood Cliffs, NJ, 1987.